

# Universal Bound on Sampling Bosons in Linear Optics

Man-Hong Yung\*

*Department of Physics, South University of Science and Technology of China, Shenzhen, Guangdong 518055, China*

Xun Gao

*Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China*

Joonsuk Huh†

*Mueunjae Institute for Chemistry (MIC), Department of Chemistry,  
Pohang University of Science and Technology (POSTECH), Pohang 790-784, Korea*

In linear optics, photons are scattered in a network through passive optical elements including beamsplitters and phase shifters, leading to many intriguing applications in physics, such as Mach-Zehnder interferometry, Hong-Ou-Mandel effect, and tests of fundamental quantum mechanics. Here we present a general analytic expression governing the upper limit of the transition amplitudes in sampling bosons, through all realizable linear optics. Apart from boson sampling, this transition bound results in many other interesting applications, including behaviors of Bose-Einstein Condensates (BEC) in optical networks, counterparts of Hong-Ou-Mandel effects for multiple photons, and approximating permanents of matrices. Also, this general bound implies the existence of a polynomial-time randomized algorithm for estimating transition amplitudes of bosons, which represents a solution to an open problem raised by Aaronson and Hance in 2012.

**Introduction.** Apart from being of fundamental interest in physics, linear optics has become a simple but powerful tool for processing quantum information [1–4] and quantum simulation [5–8]. One of the major advantages for encoding information with light is that photons are highly robust against decoherence, which makes it an ideal system to study quantum coherence [9–12]. Furthermore, in linear optical networks, all possible transformations can be achieved with simple operations involving at most a pair of modes; more precisely, every optical circuit can be implemented with beamsplitters and phase shifters only [13]. Linear optical networks have been routinely built in photonic chips [14–16] using standard semiconductor fabrication technology. In particular, a reprogrammable linear optical circuit has been integrated into a photonic chip [16], which can perform universal operations on six photonic modes with up to six photons.

Recently, it was found that boson sampling [17], as a novel application of linear optics, can be regarded as evidence for proving the inefficiency of classical devices to perform quantum simulation, which represents a serious challenge to the validity of the extended Church-Turing thesis [1]. In boson sampling [17], a product of single-photon states is injected into a linear photonic network that encodes an instance of complex matrices. The ability of simulating boson sampling with classical devices implies the ability to approximate the corresponding permanents of matrices with a multiplicative error, which is widely believed to be impossible, based on computational complexity assumptions [17]. With this motivation, much progress [16, 18–23] has been made in the experimental realization of boson samplers using linear optical devices.

Here we present a theoretical bound on the transition amplitudes for sampling bosons, from a product of Fock states to another, through linear optics. This bound limits the efficiency of sampling bosons for all possible linear optical

networks, including the behaviors of Bose-Einstein Condensates (BEC) in optical networks, and the counterparts of Hong-Ou-Mandel effects for multiple photons.

Furthermore, our bound is important for our proof on the existence of a polynomial-time randomized algorithm for approximating permanents of matrices; this result resolves an open problem [24] raised by Aaronson and Hance.

The open question [24] is: “can we estimate any linear-optical amplitude (see Eq. (6)) to  $\pm 1/\text{poly}(n)$  additive error (or better) in polynomial time?”. For the special case where the initial state is the same as that in Boson Sampling, i.e.,  $|t_1 t_2 \dots t_m\rangle = |111\dots 00\rangle$ . Aaronson and Hance confirmed that such algorithm exists through a modification of the Gurvits algorithm.

With our bound shown in Eq. (2), we confirm that there exists a polynomial-time randomized algorithm for the general cases. Thus, the open problem is now completely settled.

**Transition amplitude in linear optics.** The problem of interest in this work is described as follows: let us suppose that we are given a product of Fock states containing a total of  $n$  identical photons (or generally bosons) distributed over  $m$  different modes, i.e.,  $|t_1 t_2 \dots t_m\rangle \equiv |t_1\rangle \otimes |t_2\rangle \otimes \dots \otimes |t_m\rangle$ , where  $|t_k\rangle \equiv (t_k!)^{-1/2} (a_k^\dagger)^{t_k} |\text{vac}\rangle$  contains  $t_k$  photons for  $t_k = 0, 1, 2, \dots, n$ .

Moreover, the state is subject to the constraint of particle conservation:  $\sum_{k=1}^m t_k = n$ . Here  $a_k^\dagger$  creates a boson in  $k$ -th mode and satisfies the commutation relations:  $[a_j, a_k^\dagger] \equiv a_j a_k^\dagger - a_k^\dagger a_j = \delta_{ij}$ ,  $[a_j^\dagger, a_k^\dagger] = [a_j, a_k] = 0$ .

Let us consider any member  $U$  in the set of all possible unitary operators (i.e., linear optics) that induces a linear transformation (i.e., non-interacting) for the boson modes, i.e.,

$$U a_k^\dagger U^\dagger = \sum_{j=1}^m u_{kj} a_j^\dagger. \quad (1)$$

The central problem is to give an upper bound for the absolute value of the transition amplitude,  $|\langle s_1 s_2 \dots s_m | U | t_1 t_2 \dots t_m \rangle|$ , for locating the resulting state in another given product state,  $|s_1 s_2 \dots s_m\rangle$ , subject to the same particle-conserving constraint  $\sum_{k=1}^m s_k = n$ , for  $s_k = 0, 1, 2, \dots, n$ .

**Main result.** Here we shall prove that the upper bound of the boson transition amplitude is given by the following expression:

$$|\langle s_1 \dots s_m | U | t_1 \dots t_m \rangle| \leq \min \{v_s/v_t, v_t/v_s\}, \quad (2)$$

where  $v_s$  is a product of  $m$  factors generated from the elements in the list  $\mathbf{s} = (s_1, s_2, \dots, s_m)$ ,

$$v_s \equiv \sqrt{(s_1!/s_1^{s_1})(s_2!/s_2^{s_2}) \dots (s_m!/s_m^{s_m})}, \quad (3)$$

and defined similarly for  $v_t$ . If one of the modes is unoccupied, e.g.,  $s_k = 0$ , then we simply set  $s_k!/s_k^{s_k} \rightarrow 1$ .

In the context of boson sampling [17], the initial state is always a product of single-photon states, i.e.,  $|t_1 t_2 \dots t_m\rangle = |111 \dots 00\rangle$ . In this case, we can recover the result obtained previously by Aaronson and Hance [24], i.e.,  $|\langle s_1 s_2 \dots s_m | U | 111 \dots 00 \rangle| \leq v_s$ .

**Direct applications.** As an application, the bound found by Aaronson and Hance implies that [24] for the case where  $s_1 = n$ , and  $s_2 = s_3 \dots = s_m = 0$ , the probability of putting all bosons into the same mode from  $|111 \dots 00\rangle$  is exponentially low, as

$$P_{\max}(n, 0 \dots 0 | 1, 1 \dots 1) = v_s^2 = n!/n^n \approx \sqrt{2\pi n} e^{-n}, \quad (4)$$

using the Stirling approximation,  $n! \approx \sqrt{2\pi n} (n/e)^n$ .

Consequently, for  $n \geq 3$ , one cannot observe the counterpart of Hong-Ou-Mandel dip with linear optics [24]; the reason why Hong-Ou-Mandel dip is possible for the case of  $n = 2$  is because the bound is given by  $P_{\max}(2, 0 | 1, 1) = v_s^2 = 2!/2^2 = 1/2$ , but there are two modes; so the total probability can reach unity.

With the more general bound shown in Eq. (2), we can bound the transition probabilities for more scenarios. For example, imagine there are  $p$  bosons in one mode and  $q$  bosons in another mode. The probability of getting  $p + q$  in a single mode through linear optics is then bounded by the following,

$$P_{\max}(p + q, 0 | p, q) = \frac{(p + q)!}{p!q!} \frac{p^p q^q}{(p + q)^{p+q}}, \quad (5)$$

or its inverse. We can, for instance, ask the following questions.

**Question 1:** *can we apply linear optics to create a mode with  $2n$  bosons from two separate modes with  $n$  bosons each?* Unlikely. In this case,  $P_{\max}(2n | n, n) = 2n!/n!^2 2^{2n}$ . In the limit of Bose-Einstein Condensate (BEC), where  $n \gg 1$ , the probability bound,  $P_{\max}(2n | n, n) \approx 1/\sqrt{\pi n}$ , decreases as  $O(1/\sqrt{n})$ . An optimal strategy for achieving the bound is to apply the 50:50 beamsplitter, i.e.,  $a_1^\dagger \rightarrow (a_1^\dagger + a_2^\dagger)/\sqrt{2}$  and  $a_2^\dagger \rightarrow (a_1^\dagger - a_2^\dagger)/\sqrt{2}$ .

**Question 2:** *can we add 1 extra boson to a BEC using linear optics?* Yes! Suppose  $p = n$  and  $q = 1$ , the bound is given by  $P_{\max}(n + 1, 0 | n, 1) = (n/(n + 1))^n$ , which approaches a constant limit when  $n \rightarrow \infty$ . In fact, this bound can be saturated by the following transformation:  $U_n a_1^\dagger U_n^\dagger = \cos \theta_n a_1^\dagger + \sin \theta_n a_2^\dagger$ , where  $\sin^2 \theta_n = (n + 1)^{-1}$ . We provide justification on the validity of the optimal strategies for both questions in the appendix.

**Connection with permanents of matrices.** Another implication of our main result is related to permanents of matrices. The transition amplitude in Eq. (2) is known (see e.g. [17]) to be related to a permanent of a matrix regarding the unitary operator  $U$ :

$$\langle s_1 s_2 \dots s_m | U | t_1 t_2 \dots t_m \rangle = \frac{\text{Perm}(U_{\mathbf{s}, \mathbf{t}})}{\sqrt{s_1! \dots s_m! \cdot t_1! \dots t_m!}}, \quad (6)$$

where  $U_{\mathbf{s}, \mathbf{t}}$  is a  $n \times n$  matrix constructed by the transformation elements  $u_{kj}$  (see Eq. (1)) of the unitary operator  $U$  in the following way: create  $s_k$  copies of a row vectors  $\boldsymbol{\mu}_{k, \mathbf{t}}$  that contains  $t_j$  copies of  $u_{kj}$ 's. For example, if  $\mathbf{s} = (1, 0, 2)$ , and  $\mathbf{t} = (2, 1, 0)$ , then the matrix  $U_{\mathbf{s}, \mathbf{t}}$  is of the following form:

$$U_{\mathbf{s}, \mathbf{t}} = \begin{bmatrix} \boldsymbol{\mu}_{1, \mathbf{t}} \\ \boldsymbol{\mu}_{3, \mathbf{t}} \\ \boldsymbol{\mu}_{3, \mathbf{t}} \end{bmatrix} = \begin{bmatrix} u_{1,1} & u_{1,1} & u_{1,2} \\ u_{3,1} & u_{3,1} & u_{3,2} \\ u_{3,1} & u_{3,1} & u_{3,2} \end{bmatrix}. \quad (7)$$

Note that if all  $s$ 's and  $t$ 's equal unity, then the transition probability is exactly the same as the permanent of the matrix defined in Eq. (1), i.e.,  $\langle 11 \dots 1 | U | 11 \dots 1 \rangle = \text{Perm}(u_{kj})$ . Therefore, our bound also implies an upper bound of the permanent of a matrix:

$$|\text{Perm}(U_{\mathbf{s}, \mathbf{t}})| \leq \min \left\{ \frac{v_s}{v_t}, \frac{v_t}{v_s} \right\} \times \prod_{k=1}^m \sqrt{s_k! t_k!}. \quad (8)$$

Before we present the proof of the bound, let us first establish a general theorem that is crucial for our result:

**Theorem 1.** Given any polynomial function,  $f(a_1^\dagger, a_2^\dagger, \dots, a_m^\dagger)$ , of multi-mode creation operators  $a_k^\dagger$ 's, the vacuum-to-vacuum transition amplitude (unnormalized),

$$F_s \equiv \langle \text{vac} | a_1^{s_1} a_2^{s_2} \dots a_m^{s_m} f(a_1^\dagger, a_2^\dagger, \dots, a_m^\dagger) | \text{vac} \rangle, \quad (9)$$

can always be expressed as a sum involving a set of weighted complex roots of unity, by mapping the boson operator,  $a_k^\dagger \rightarrow z_k$ , to a complex number  $z_k$ , and similarly  $a_k \rightarrow \bar{z}_k$  to its complex conjugate  $\bar{z}_k$ :

$$F_s = \frac{v_s^2}{d^m} \sum_{\{z\}} \bar{z}_1^{s_1} \bar{z}_2^{s_2} \dots \bar{z}_m^{s_m} f(z_1, z_2, \dots, z_m), \quad (10)$$

where  $z_j \in \{\sqrt{s_j} \omega^0, \sqrt{s_j} \omega^1, \dots, \sqrt{s_j} \omega^{d-1}\}$  is related to one of the complex roots of unity  $\omega \equiv e^{-2\pi i/d}$ , weighted by a factor  $\sqrt{s_j}$ . Here  $d$  is chosen to be an integer larger than the degree of the function and the sum  $\sum_{k=1}^m s_k$ .

Alternatively, we can write  $F_s$  in the form of an expectation value:

$$F_s = v_s^2 \mathbb{E}[\bar{z}_1^{s_1} \bar{z}_2^{s_2} \cdots \bar{z}_m^{s_m} f(z_1, z_2, \dots, z_m)], \quad (11)$$

which allows us to devise a sampling algorithm to estimate its value, as we shall discuss later.

**Proof of Theorem 1.** Since all the terms in the function  $f(a_1^\dagger, a_2^\dagger, \dots, a_m^\dagger)$  commute with one another, we can, for example, sort out the first creation operator  $a_1^\dagger$  as if it was just a real number, and write

$$f(a_1^\dagger, a_2^\dagger, \dots, a_m^\dagger) = \sum_{k=0}^d a_1^{\dagger k} \phi_k(a_2^\dagger, \dots, a_m^\dagger), \quad (12)$$

where  $\phi_k(a_2^\dagger, \dots, a_m^\dagger)$  is a resulting polynomial function without  $a_1^\dagger$ . Consequently, we have  $F_s = \sum_{k=0}^d \langle 0 | a_1^{s_1} a_1^{\dagger k} | 0 \rangle \langle \text{vac} | a_2^{s_2} \cdots a_m^{s_m} \phi_k(a_2^\dagger, \dots, a_m^\dagger) | \text{vac} \rangle$ , but there is only one non-zero term in the summation, as  $\langle 0 | a_1^{s_1} a_1^{\dagger k} | 0 \rangle = s_1! \delta_{s_1 k}$ .

Now, since the Kronecker delta function can be expressed (by the representation through discrete Fourier transform) as follows:  $\delta_{s_1 k} = (1/d) \sum_{j=0}^{d-1} e^{-(2\pi i j/d)(s_1 - k)} = (1/d) \sum_{j=0}^{d-1} \omega^{j(s_1 - k)}$ , we can therefore write the inner product (with  $z_1 \in \{\sqrt{s_1} \omega^0, \sqrt{s_1} \omega^1, \dots, \sqrt{s_1} \omega^{d-1}\}$ ),

$$\langle 0 | a_1^{s_1} a_1^{\dagger k} | 0 \rangle = \frac{s_1!}{s_1!} \frac{1}{d} \sum_{\{z_1\}} \bar{z}_1^{s_1} z_1^k, \quad (13)$$

as a sum over all values of  $z_1$ , which implies that  $F_s = (s_1! / s_1! d) \sum_{\{z_1\}} \bar{z}_1^{s_1} \langle \text{vac} | a_2^{s_2} \cdots a_m^{s_m} f(z_1, a_2^\dagger, \dots, a_m^\dagger) | \text{vac} \rangle$ . Next, we can define a new polynomial function,  $f'(a_2^\dagger, \dots, a_m^\dagger) \equiv (s_1! / s_1! d) \sum_{\{z_1\}} \bar{z}_1^{s_1} f(z_1, a_2^\dagger, \dots, a_m^\dagger)$ , and repeat the same procedure for  $a_2^\dagger$ , and so on, which yields the result in Eq. (10) at the end.  $\square$

**Proving the main result.** We are now ready to present the proof for the bound in Eq. (2). For this purpose, we express the transition amplitude explicitly with bosonic operators, i.e.,

$$\langle s_1 \dots s_m | U | t_1 \dots t_m \rangle = \frac{G(U, \mathbf{s}, \mathbf{t})}{\sqrt{(s_1! \cdots s_m!)(t_1! \cdots t_m!)}}, \quad (14)$$

where we defined an operator function,

$$G(U, \mathbf{s}, \mathbf{t}) \equiv \langle \text{vac} | a_1^{s_1} \cdots a_m^{s_m} U a_1^{\dagger t_1} \cdots a_m^{\dagger t_m} | \text{vac} \rangle. \quad (15)$$

The proof can be completed with only three steps as follows.

**Step 1 (operator-to-number conversion):** With the transformation rule given in Eq. (1), we have  $U a_1^{\dagger t_1} \cdots a_m^{\dagger t_m} U^\dagger = \prod_{k=1}^m (u_{k,1} a_1^\dagger + \dots + u_{k,m} a_m^\dagger)^{t_k}$ , which is exactly a polynomial function of the creation operators. Therefore, the theorem above implies that

$$G(U, \mathbf{s}, \mathbf{t}) = \frac{v_s^2}{d^m} \sum_{\{\mathbf{z}\}} (\bar{z}_1^{s_1} \bar{z}_2^{s_2} \cdots \bar{z}_m^{s_m}) g(\mathbf{z}), \quad (16)$$

where the function  $g(\mathbf{z})$  is defined as follows:

$$g(\mathbf{z}) \equiv \prod_{k=1}^m (u_{k,1} z_1 + \dots + u_{k,m} z_m)^{t_k}. \quad (17)$$

In order to bound the absolute value of  $G(U, \mathbf{s}, \mathbf{t})$ , it is sufficient to bound the function  $g(\mathbf{z})$  by writing its absolute value in the following form:  $|g(\mathbf{z})| = \sqrt{t_1^{t_1} \cdots t_m^{t_m}} \prod_{j=1}^m (1/\sqrt{t_j})^{t_j} |u_{k,1} z_1 + \dots + u_{k,m} z_m|^{t_j}$ .

**Step 2 (arithmetic-geometric inequality):** Recall that the weighted arithmetic-geometric inequality suggests that,

$$A_1^{\lambda_1} A_2^{\lambda_2} \cdots A_m^{\lambda_m} \leq \lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_m A_m, \quad (18)$$

for all non-negative  $A_k$ 's and  $\lambda_k$ , subject to the constraint  $\sum_{k=1}^m \lambda_k = 1$ . In terms of our  $t$ 's (by setting  $\lambda_k = t_k/n$ ), we have  $(A_1^{t_1} A_2^{t_2} \cdots A_m^{t_m})^{1/2} \leq [(t_1/n) A_1 + (t_2/n) A_2 + \dots + (t_m/n) A_m]^{n/2}$ . Now, let us denote  $A_j = (1/t_j) |u_{k,1} z_1 + \dots + u_{k,m} z_m|^2$ . Then, we have,

$$\frac{|g(\mathbf{z})|}{\sqrt{t_1^{t_1} \cdots t_m^{t_m}}} \leq \left( \sum_{j=1}^m \frac{1}{n} |u_{k,1} z_1 + \dots + u_{k,m} z_m|^2 \right)^{n/2}. \quad (19)$$

**Step 3 (bounding the norms):** Note that the right-hand side is related to the 2-norm of a vector:  $\|\mathbf{z}\| \equiv \|\mathbf{z}\|_2 = \sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_m|^2}$ . To take a step further, we can always define a unitary matrix  $V$  such that  $(V\mathbf{z})_k = u_{k,1} z_1 + \dots + u_{k,m} z_m$ , which implies that  $|g(\mathbf{z})| \leq (t_1^{t_1} \cdots t_m^{t_m} / n^n)^{1/2} \|V\mathbf{z}\|^{n/2}$ . Since  $\|V\| = 1$  for unitary matrices, and  $\|\mathbf{z}\| = (s_1 + s_2 + \dots + s_m)^{1/2} = \sqrt{n}$ , we have  $|g(\mathbf{z})| \leq (t_1^{t_1} \cdots t_m^{t_m})^{1/2}$ . Consequently, we have

$$|G(U, \mathbf{s}, \mathbf{t})| \leq v_s^2 \sqrt{(s_1^{s_1} \cdots s_m^{s_m})(t_1^{t_1} \cdots t_m^{t_m})}, \quad (20)$$

which implies part of the advertised result of the bound  $v_s/v_t$  in Eq. (2). We can repeat essentially the same procedure for the complex conjugate,  $\langle t_1 \cdots t_m | U^\dagger | s_1 \cdots s_m \rangle$ , of the transition amplitude, in order to obtain the other part,  $v_s/v_t$ . This completes our proof.  $\blacksquare$

**Permanent by Sampling.** It is known that any  $m \times m$  matrix  $W = (w_{i,j})$  permanent can be calculated exactly with a scaling  $O(m^2 2^m)$  using Ryser's formula [25]. Glynn [26, 27] suggested a different algorithm requiring a similar computational cost that the Glynn's formula is given as a normalized form:  $\text{Perm}(W) = 2^{-m} \sum_{\mathbf{x}} x_1 \cdots x_m \prod_{i=1}^m (w_{i,1} x_1 + \dots + w_{i,m} x_m)$ , summing over all possible  $m$ -bit strings  $\mathbf{x} = (x_1, x_2, \dots, x_m) \in \{\pm 1\}^m$ , then normalizing with the total number of summations  $2^m$ . Based on the Glynn's formula, Gurvits [28] proposed a polynomial-time randomized algorithm to produce an approximation  $\text{Perm}(W)$  to the value of the permanent, with an additive error  $\pm \epsilon \|W\|^m$ , i.e.,  $|\tilde{\text{P}}(W) - \text{Perm}(W)| \leq \epsilon \|W\|^m$ , where  $\|W\| \equiv \sup_{\mathbf{v} \neq 0} \|W\mathbf{v}\| / \|\mathbf{v}\|$ .

The main idea of Gurvits is that one can convert Glynn's formula into an expectation value:

$\text{Perm}(W) = \mathbb{E}[x_1 \cdots x_m \prod_{i=1}^m (w_{i,1}x_1 + \dots + w_{i,n}x_m)]$ , where  $\text{Gly}(x) \equiv x_1 \cdots x_m \prod_{i=1}^m (w_{i,1}x_1 + \dots + w_{i,m}x_m)$  is Glynn's estimator [26]. An approximation of the permanent is obtained by randomly and uniformly picking  $T$  strings  $x_k \in \{\pm 1\}^m$ , for  $k = 1, 2, \dots, T$ , and evaluate the average value:

$$\tilde{P}(W) = \frac{1}{T} \sum_{k=1}^T \text{Gly}(x_k). \quad (21)$$

As a result, if we take a total of  $T = O(m^2/\epsilon^2)$  samples, then the Chebyshev's inequality guarantees that the failure probability, where  $|\tilde{P}(W) - \text{Perm}(W)| > \epsilon \|W\|^m$ , can be upper-bounded with a small value.

In Ref. [24], Aaronson and Hance proposed a generalization of Gurvits's algorithm by defining a generalized Glynn's estimator, namely  $\text{GenGly}(z) \equiv v_s^2 (\tilde{z}_1^{s_1} \cdots \tilde{z}_m^{s_m}) \prod_{i=1}^m (w_{i,1}z_1 + \dots + w_{i,m}z_m)$ , where  $v_s$  is defined in Eq. (3). Sampling the generalized Glynn's estimator over the complex values, the permanent,  $\text{Perm}(V)$ , of a matrix  $V$ , which is obtained by repeating  $s_i$  times the  $i$ -th row of the  $m \times m$  matrix  $W = (w_{i,j})$ , can be estimated in polynomial time with an additive error  $\pm \epsilon v_s \sqrt{s_1! \cdots s_m!} \|W\|^n$ .

**Sampling algorithm for transition amplitudes.** Comparing the right-hand sides of Eq. (6) and Eq. (14), we concluded that  $G(U, s, t)$  is equal to the permanent of the matrix  $U_{s,t}$ , i.e.,

$$G(U, s, t) = \text{Perm}(U_{s,t}). \quad (22)$$

It is therefore possible to extend our formalism for an arbitrary  $m \times m$  matrix  $W = (w_{i,j})$  from the transformation in Eq. (1), which implies that we can define an even more general Glynn estimator,

$$\text{mGenGly}(z) \equiv v_s^2 \left( \prod_{k=1}^m \tilde{z}_k^{s_k} \right) \prod_{i=1}^m \left( \sum_{j=1}^m w_{i,j} z_j \right)^{t_i}, \quad (23)$$

which is reduced to the estimator,  $\text{GenGly}(z)$ , of Aaronson and Hance for the cases where  $t_1 = t_2 = \dots = t_m = 1$ , and further reduced to the estimator,  $\text{Gly}(z)$ , of Gurvits, when  $s_1 = s_2 = \dots = s_m = 1$  in addition. An alternative estimator can be found in Huh [29].

As a result, taking a total of  $T = O(1/\epsilon^2)$  samples, the error in estimating the permanent is bounded by  $\pm \epsilon \|W\|^m v_s^2 \prod_{k=1}^m (s_k^{s_k} t_k^{t_k})^{1/2}$ , derived from the bound established in Eq. (20). Note that the evaluation of each sample requires  $O(m^2)$  steps, as in Eq. (23), the calculation of the summation takes  $m$  steps and there are  $m$  factors to multiply.

Return to the case of quantum optics, where the transformation is necessarily a unitary matrix  $U$ , with  $\|U\| = 1$ . With the identification in Eq. (22) and our bound in Eq. (2), we can therefore approximate the transition amplitude with a high probability, by uniformly sampling the

more general Glynn's estimator in Eq. (23), with  $z_k \in \{\sqrt{s_j} \omega^0, \sqrt{s_j} \omega^1, \dots, \sqrt{s_j} \omega^{d-1}\}^m$ , i.e.,

$$\langle s_1 \cdots s_m | U | t_1 \cdots t_m \rangle \approx \frac{1}{T} \frac{\sum_{k=1}^T \text{mGenGly}(z_k)}{\prod_{k=1}^m (s_k! t_k!)^{1/2}}. \quad (24)$$

With a polynomial number of sampling,  $T = O(1/\epsilon^2)$ , the error of the approximation of the transition amplitude in Eq. (24) is bounded by  $\pm \epsilon \times \min\{v_s/v_t, v_t/v_s\}$ , from the Chebyshev's inequality (see Appendix for details). The existence of this polynomial-time algorithm, scaling as  $O(m^2/\epsilon^2)$ , represents a solution to an open problem raised in the work [24] of Aaronson and Hance.

**Conclusion.** We have presented a general upper bound (Eq. (2)) on the transition amplitudes in sampling bosons for any linear optical network (Eq. (1)). This bound can directly be applied to many different physical scenarios such as Hong-Ou-Mandel and BEC (see Eq. (4) and Eq. (5), respectively). The crucial step in proving the bound involves a general theorem (see Eq. (9)) that makes it possible to convert any vacuum-to-vacuum transition amplitude, for some polynomial functions of the boson operators, into a sum of discrete random variables (Eq. (10)). In addition to boson sampling, this theorem is applicable to the calculation of elements of the  $S$ -matrix in quantum electrodynamics (see Ref. [30]).

The connection between the transition amplitudes and the permanents makes it possible to bound the absolute value of the corresponding permanents of matrices (Eq. (8)). Moreover, the classical algorithm proposed by Gurvits [28], extended by Aaronson and Hance [24], can be further extended (Eq. (24)) with our bound; the existence of such algorithm implies that the open problem of Aaronson and Hance in Ref. [24] (page 16) can now be settled.

Finally, we note that it is straightforward to show that our bound can also be applied to generalize the de-randomizing algorithm for approximating permanents of non-negative matrices, which was discussed by Aaronson and Hance [24].

**Acknowledgments**— We thank Scott Aaronson for the valuable comments on the manuscript. M.-H.Y. acknowledges support by the National Natural Science Foundation of China under Grants No. 11405093. J.H. acknowledges supports by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(NRF-2015R1A6A3A04059773), the ICT R&D program of MSIP/IITP [2015-019, Fundamental Research Toward Secure Quantum Communication] and Mueunjae Institute for Chemistry (MIC) postdoctoral fellowship.

\* yung@sustc.edu.cn  
† huhsh@postech.ac.kr

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*.



- Quantum Computation and Quantum Information (Cambridge University Press, 2011).
- [2] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, *Rev. Mod. Phys.* **79**, 135 (2007).
- [3] A. Peruzzo, J. McClean, P. Shadbolt, M.-h. Yung, X.-q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *Nat. Commun.* **5**, 4213 (2014).
- [4] R. J. Chapman, M. Santandrea, Z. Huang, G. Corrielli, A. Crespi, M.-H. Yung, R. Osellame, and A. Peruzzo, *Nat. Commun.* **7**, 11339 (2016).
- [5] C.-y. Lu, W.-b. Gao, O. Gühne, X.-q. Zhou, Z.-b. Chen, and J.-w. Pan, *Phys. Rev. Lett.* **102**, 030502 (2009).
- [6] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, *Nat. Chem.* **2**, 106 (2010).
- [7] A. Aspuru-Guzik and P. Walther, *Nat. Phys.* **8**, 285 (2012).
- [8] J.-S. Xu, M.-H. Yung, X.-y. Xu, S. Boixo, Z.-w. Zhou, C.-f. Li, A. Aspuru-Guzik, and G.-C. Guo, *Nat. Photonics* **8**, 113 (2014).
- [9] M. O. Scully and M. S. Zubairy, *Quantum optics* (Cambridge university press, 1997).
- [10] J.-S. Xu, M.-H. Yung, X.-Y. Xu, J.-S. Tang, C.-F. Li, and G.-C. Guo, *Sci. Adv.* **2**, e1500672 (2016).
- [11] M. Tillmann, S.-H. Tan, S. E. Stoeckl, B. C. Sanders, H. de Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Phys. Rev. X* **5**, 041015 (2015).
- [12] D. N. Biggerstaff, R. Heilmann, A. a. Zezevik, M. Gräfe, M. a. Broome, A. Fedrizzi, S. Nolte, A. Szameit, A. G. White, and I. Kassal, *Nat. Commun.* **7**, 11282 (2016).
- [13] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Phys. Rev. Lett.* **73**, 58 (1994).
- [14] A. Crespi, R. Ramponi, R. Osellame, L. Sansoni, I. Bongioanni, F. Sciarrino, G. Vallone, and P. Mataloni, *Nat. Commun.* **2**, 566 (2011).
- [15] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. F. Matthews, M. G. Thompson, and J. L. O'Brien, *Nat. Photonics* **6**, 45 (2011).
- [16] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, *Science* **349**, 711 (2015).
- [17] S. Aaronson and A. Arkhipov, *Theory Comput.* **9**, 143 (2013).
- [18] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, *Science* (80-. ) **339**, 794 (2013).
- [19] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley, *Science* **339**, 798 (2013).
- [20] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Nat. Photonics* **7**, 540 (2013).
- [21] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, *Nat. Photonics* **7**, 545 (2013).
- [22] N. Spagnolo, C. Vitelli, M. Bentivegna, D. J. Brod, A. Crespi, F. Flamini, S. Giacomini, G. Milani, R. Ramponi, P. Mataloni, R. Osellame, E. F. Galvão, and F. Sciarrino, *Nat. Photonics* **8**, 615 (2014).
- [23] M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Vigi-gianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvao, A. Crespi, R. Ramponi, R. Osellame, and F. Sciarrino, *Sci. Adv.* **1**, e1400255 (2015).
- [24] S. Aaronson and T. Hance, *Quantum Inf. Comput.* **14**, 541 (2012).
- [25] H. J. Ryser, *Combinatorial mathematics* (Mathematical Association of America, 1963).
- [26] D. G. Glynn, *Eur. J. Comb.* **31**, 1887 (2010).
- [27] D. G. Glynn, *Designs, Codes, and Cryptography* **68**, 39 (2013).
- [28] L. Gurvits, in *Math. Found. Comput. Sci.* **2005** (2005) pp. 447–458.
- [29] J. Huh, arXiv:1605.08506v1 (2016).
- [30] E. R. Caianiello, *Nuovo Cim.* **10**, 1634 (1953).

## Appendix: Validity of the optimal strategies

Here we show that the upper bounds of the two examples can be saturated with the strategies discussed in the main text. First, consider putting  $n$  bosons from two modes,  $a_1^\dagger$  and  $a_2^\dagger$ , into  $2n$  bosons in one mode. For a rotational operator  $U$ , where

$$U a_1^\dagger U^\dagger = \cos \theta a_1^\dagger + \sin \theta a_2^\dagger, \quad (25)$$

$$U a_2^\dagger U^\dagger = \cos \theta a_2^\dagger - \sin \theta a_1^\dagger. \quad (26)$$

The corresponding transition amplitude  $A(2n, 0|n, n)$  is given by,

$$A(2n, 0|n, n) = \frac{\langle \text{vac} | a_1^{2n} U a_1^\dagger a_2^\dagger | \text{vac} \rangle}{\sqrt{(2n)!n!n!}}, \quad (27)$$

where  $\langle \text{vac} | a_1^{2n} U a_1^\dagger a_2^\dagger | \text{vac} \rangle = \sin^n(2\theta) (2n)!/2^n$ . Therefore, the maximum probability can be achieved by setting  $\theta = \pi/4$  (i.e., a 50:50 beamsplitter), which coincides with the upper bound in the main text,

$$P_{\max}(2n, 0|n, n) = |A|^2 = \frac{(2n)!}{(n!)^2 2^{2n}}. \quad (28)$$

For the second question, where an extra boson is added to a mode with  $n$  bosons, we choose the same form of  $U$  as in Eq. (25) and Eq. (26), but with  $\sin^2 \theta_n = (n+1)^{-1}$ . In this way, we have

$$A(n+1, 0|n, 1) = \frac{\langle \text{vac} | a_1^{n+1} U a_1^\dagger a_2^\dagger | \text{vac} \rangle}{\sqrt{(n+1)!n!}}, \quad (29)$$

where  $\langle \text{vac} | a_1^{n+1} U a_1^\dagger a_2^\dagger | \text{vac} \rangle = \cos^n \theta \sin \theta (n+1)!$ . Therefore, the maximum probability is given by

$$P_{\max}(n+1, 0|n, 1) = \left( \frac{n}{n+1} \right)^n, \quad (30)$$

which coincides with the upper bound.

## Appendix: Chebyshev's inequality for random variables with discrete complex values

Let us discuss more on the Chebyshev's inequality required for our bound. The majority of textbooks deals with discrete

real random variables, but here we are considering discrete complex numbers. To ensure the inequality is still applicable to our case, we provide an appendix to derive it by assuming complex random variables in the beginning.

Consider the generalization of the Markov inequality for a discrete random variable  $X$ , which takes on discrete complex values from a set  $\{x_i\}$ . The Markov's bound implies that, for a non-negative  $\alpha$ ,

$$\Pr[|X| \geq \alpha] \leq \frac{\mathbb{E}(|X|)}{\alpha}. \quad (31)$$

---

*Proof.* By definition, the expectation value of the absolute values is given by  $\mathbb{E}(|X|) = \sum_a a \cdot \Pr(|X| = a)$ , where the sum is over all possible values of  $|X|$ , labeled by  $a$ . We can sort out the parts where  $a \geq \alpha$ , which means that,

$$\mathbb{E}(|X|) \geq \sum_{a \geq \alpha} a \cdot \Pr(|X| = a). \quad (32)$$

Furthermore, we also get a smaller value by replacing all  $a$ 's by  $\alpha$ , i.e.,

$$\mathbb{E}(|X|) \geq \alpha \sum_{a \geq \alpha} \Pr(|X| = a). \quad (33)$$

Lastly, we can identify the last term as the probability where  $|X| \geq \alpha$ , i.e.,

$$\Pr[|X| \geq \alpha] = \sum_{a \geq \alpha} \Pr(|X| = a), \quad (34)$$

which yields the desired result.  $\square$

---

Now, define a variable,  $\mu$ , to represent the expectation value of  $X$ , i.e.,

$$\mu \equiv \mathbb{E}(X) = \sum_i x_i \cdot \Pr(X = x_i). \quad (35)$$

Furthermore, we define a new random variable

$$Y \equiv |X - \mu|^2 = (X - \mu)(X^* - \mu^*), \quad (36)$$

where  $\mathbb{E}(Y) = \sum_i |x_i - \mu|^2 \Pr(X = x_i)$ . Note that we can also write,

$$\mathbb{E}(Y) = \text{Var}(X) \equiv \mathbb{E}(|X|^2) - |\mu|^2, \quad (37)$$

and that,  $\Pr(|X - \mu| \geq \alpha) = \Pr(Y \geq \alpha^2)$ . Therefore, if we apply the Markov inequality as follows,

$$\Pr(Y \geq \alpha^2) \leq \frac{\mathbb{E}(Y)}{\alpha^2} = \frac{\text{Var}(X)}{\alpha^2}, \quad (38)$$

then, we obtain the Chebyshev's inequality,

$$\Pr(|X - \mu| \geq \alpha) \leq \frac{\text{Var}(X)}{\alpha^2}. \quad (39)$$

Now, let us consider  $n$  independent complex random variables,  $X_1, X_2, \dots, X_n$ , with  $\mathbb{E}(X_i) = \mu_i$  and  $\text{Var}(X_i) = \sigma_i^2$ . Then, the Chebyshev's inequality implies that

$$\Pr\left(\left|\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i\right| \geq \alpha\right) \leq \frac{\sum_{i=1}^n \sigma_i^2}{\alpha^2}, \quad (40)$$

where we used the fact that  $\text{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \text{Var}(X_i)$ .

For identical random variables, where all  $\mu_i = \mu$  and  $\sigma_i = \sigma$ , we have

$$\Pr\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}, \quad (41)$$

which is the inequality needed for our results.